Minimally Incomplete Sampling and Convergence of Adaptive Play in 2×2 Games

Ethan Holdahl* Anne van den Nouweland†

June 13, 2023

Abstract

Keywords:

JEL Classification Codes:

*University of Oregon

†Department of Economics, University of Oregon, Eugene, OR 97403-1285. Email: annev@uoregon.edu.

1 Introduction

With over 1000 citations, Young [1993] is a seminal paper in the field of evolutionary game theory. In it, Young introduces a model of learning called adaptive play in which players best respond to a sampled history of play. Young proved that play will eventually converge to a convention, a self-enforcing pattern of play where the same Nash equilibrium is played in each period, if the sampling in the available history by the players is sufficiently incomplete. Through this backward looking best response behavior, Young offers an explanation for how order and norms can spontaneously evolve in populations.

In the process of adaptive play, occasionally players make mistakes and play an action that is not a best response. Such instances are called perturbations and allow the perturbed adaptive play process to escape a convention and travel to another one. The resistance of moving from one convention to another is measured by the number of mistakes that are necessary to move the process into the basin of attraction of the latter convention. The conventions that require the most mistakes to move from and/or the fewest to move to are most likely to be played in the long run. Such conventions are said to be stochastically stable. The theory of adaptive play, sometimes called adaptive learning or fictitious play with bounded memory, and its most celebrated result, identifying the stochastically stable patterns of play, have since been applied to a wide variety of games. We restrict the scope of our literature review here to games with finitely many strategies in which groups of players interact with each other on a random basis.[1]

In his book *Individual Strategy and Social Structure* (Young [1998]), Young expanded upon the foundation he laid in Young [1993]. He proved that in 2×2 coordination games adaptive play eventually converges to a convention if the amount sampled (s) in the available memory (m) is sufficiently incomplete, where the criterion for "sufficiently" incomplete is slackened to $s/m \leq 1/2$ in Young [1998] (from the more permissive restriction in Young [1993][2]). Most recently, Proposition 6.4 of Wallace and Young [2015] simply states that in n-player coordination games "if s/m is sufficiently small, the [adaptive learning] process converges with probability one to a convention from any initial state". Although acknowledged by Young in his 1998 book, "We do not claim the bound on incompleteness $s/m \leq 1/2$ is the best possible", to our knowledge no one has found and proven what the best possible bound is. Consequently, follow-up work building upon this theory has retained the restrictive bound of $s/m \leq 1/2$.

The literature we review applies adaptive learning to a variety of games with a focus on finding different

[1] So, we do not, for example, consider games played on networks. For an introduction and references to the extensive literature on such games, we refer to reader to Wallace and Young [2015].

[2] The result in Young [1993] that we are referring to is Theorem 1, which is formulated for a more general setting than only 2×2 coordination games.

criteria for stochastic stability. Maruta [1997] and Ellison [2000] independently introduced the idea of global risk dominance,[3] which indicates the existence of an action that risk dominates every other strategy in a game, and showed that the stochastic adaptive learning process of Young [1993] selects the globally risk dominant strategy if there is one.

Durieu et al. [2011] applied Young's adaptive play model to the concept of p-best response sets. The idea of p-dominant equilibrium, first introduced by Morris et al. [1995], was adapted by Tercieux [2006] to create the concept of p-best response sets. A p-best response set is a cartesian product of strategies for each player such that when all players believe all other player(s) will play a strategy contributing to the p-best response set with at least probability p then the best response(s) of all players are strategies contributing to the p-best response set. A p-best response set with $p = 1$ is therefore equivalent to the concept of a so-called curb set as introduced by Basu and Weibull [1991] and discussed in the context of adaptive play in Chapter 7 of Young [1998]. A p-best response set is said to be minimal if it contains no proper subsets which are also p-best response sets. Durieu et al. [2011] show that in n-person games, if p is sufficiently small then there is a unique minimal p-best response set and only the strategy profiles contained within the unique minimal p-best response set can be stochastically stable given perturbation rates are also sufficiently small. Note that this method does not necessarily make as sharp a prediction as Young [1993] since a p-best response set may contain multiple conventions of which not all are stochastically stable.

Breaking off from the canonical model, adaptive play has also been modified to fit a cognitive hierarchy framework [Sáez-Martí and Weibull, 1999, Matros, 2003, Khan and Peeters, 2014]. This branch of theory allows for variability in the degree of sophistication through which players compute their best response, similar to level-k thinking as introduced by Nagel [1995]. In Young's model of adaptive play, agents are backwards looking and best respond to their sample of their opponents' play. Sáez-Martí and Weibull [1999] and Matros [2003] refer to these players as "not clever" and Khan and Peeters [2014] refers to them as "level-1" individuals. One step higher on the cognitive scale are the "clever" and "level-2" individuals. These players sample their own history of play, compute their opponents' best responses to that, and then play their own best response to their opponents' predicted play. Sáez-Martí and Weibull [1999] and Matros [2003] cap the cognitive hierarchy at level-2 whereas Khan and Peeters [2014] allows higher levels and moreover has even-leveled individuals sample their own history and odd-leveled individuals sample their opponents' history. Sáez-Martí and Weibull [1999] studies bargaining games, Matros [2003] studies generic two-player games, and Khan and Peeters [2014] covers both of these. All three impose $s \leq m/2$[4] and find that introducing

[3] The terminology "global risk dominance" was introduced in Maruta [1997].

[4] The authors allow for different roles in the game to have different sample sizes but impose the upper limit of $m/2$ for all

"clever" agents can change which states are stochastically stable. Matros [2003] and Khan and Peeters [2014] find that play with "clever" agents still converges to a minimal curb set.

Jensen et al. [2005] applies adaptive play to static 2-player games of incomplete information. They allow for different types of players within the class of players for each role in the game, where each type of player creates their own history. In each period, only the memories of the types who are selected to play are updated, and those of the other types remain unchanged. Players know the distribution of types in the other classes and sample for each type from the most recent m periods in which that type played, and then weight their sample by the prevalence of each type and subsequently compute their best response. Jensen et al. [2005] examines in detail a 2×2 game of chicken with incomplete information and leverages the condition $s < m/4$ (which they obtain by applying Theorem 1 in Young [1993]) to show that the basic learning process converges to a convention and that convention may be one which "lacks coordination" where not all types for the same player play the same strategy. Depending on the payoffs in the game, this "uncoordinated" convention can be stochastically stable.

In our paper we show that *any* degree of incomplete sampling is sufficient for the unperturbed adaptive play process to converge to an equilibrium in 2×2 coordination games from any given history. In addition, we show that incomplete sampling is unnecessary in all but some 2×2 games. We also show that increasing the sample size beyond $s/m \leq 1/2$ may result in increased levels of resistance between conventions, that is to say, increasing the sample size may make conventions more stable. However, even though the resistance between conventions may change, we show that this change does not affect which convention(s) are stochastically stable when sampling is incomplete ($s < m$).

2 Adaptive play in 2×2 coordination games

2.1 2×2 coordination games

Consider a 2×2 game $G = (N; A_1, A_2; u_1, u_2)$ with player set $N = \{1, 2\}$, actions sets $A_1 = \{a_1, a_2\}$ and $A_2 = \{b_1, b_2\}$, and payoff functions $u_i : A_1 \times A_2 \to \mathbf{R}$ $(i = 1, 2)$. The game G is a coordination game if it has two pure-strategy Nash equilibria on a diagonal. Without loss of generality, we assume that (a_1, b_1) and (a_2, b_2) are Nash equilibria and we also assume that for player 1 either $u_1(a_1, b_1) > u_1(a_2, b_1)$ or $u_1(a_2, b_2) > u_1(a_1, b_2)$ and for player 2 either $u_2(a_1, b_1) > u_2(a_1, b_2)$ or $u_2(a_2, b_2) > u_2(a_2, b_1)$. These last two conditions rule out the possibility that one of the players has the same payoff from both their actions

samples.

regardless of the action played by the other player, in which case the only distinction between a player's two actions is, from their own perspective, the names of the actions.

Because each player has only two actions in the game G, every mixed strategy p_i of player $i \in \{1, 2\}$ can be identified by the probability $p_i(s_i)$ with which player i plays one of their actions s_i (because that leaves probability $1 - p_i(s_i)$ that player i plays their other action). Action a_1 is a best response by player 1 to a mixed strategy p_2 of player 2 if and only if $p_2(b_1) \geq \frac{u_1(a_2,b_2)-u_1(a_1,b_2)}{u_1(a_1,b_1)-u_1(a_1,b_2)-u_1(a_2,b_1)+u_1(a_2,b_2)}$. Note that $\alpha_2 := \frac{u_1(a_2,b_2)-u_1(a_1,b_2)}{u_1(a_1,b_1)-u_1(a_1,b_2)-u_1(a_2,b_1)+u_1(a_2,b_2)} \in [0, 1]$ because G is a coordination game with Nash equilibria (a_1, b_1) and (a_2, b_2), and either $u_1(a_1, b_1) > u_1(a_2, b_1)$ or $u_1(a_2, b_2) > u_1(a_1, b_2)$.[5] Similarly, action b_1 is a best response by player 2 to a mixed strategy p_1 of player 1 if and only if $p_1(a_1) \geq \alpha_1 := \frac{u_2(a_2,b_2)-u_2(a_2,b_1)}{u_2(a_1,b_1)-u_2(a_2,b_1)-u_2(a_1,b_2)+u_2(a_2,b_2)} \in [0, 1]$.

2.2 Adaptive play in 2-player games

We study adaptive play [Young, 1993] with memory m and sample size $1 < k \leq m$ of the game G, as explained below.[6]

For each role (player position) $i \in N$ in game G, there is a class of players C_i who can play that role. No player can play in more than one role ($C_1 \cap C_2 = \emptyset$). In each period t, a player is drawn from each class, and the two players that are drawn play the game G – each player i chooses an action $s_i(t) \in A_i$ from the actions available to them in their role. The action-tuple $s(t) = (s_1(t), s_2(t))$ is recorded and will be referred to as the play at time t. The history of plays up to and including time t is the ordered vector $h(t) = (s(1), s(2), s(3), ..., s(t))$, and the history of the last m plays, called a state, is the ordered vector $h(t|m) = (s(t - m + 1), s(t - m + 2), ..., s(t))$.

In period $t + 1$, the player in role i draws a sample R_i^{t+1} of size k from the m most recent plays $s_j(t - m + 1), s_j(t - m + 2), \ldots, s_j(t)$ by the players in role $j \neq i$. Player i predicts that the players in role j play a mixed strategy $p_j(\cdot|R_i^{t+1})$ that is the frequency distribution of the actions in the sample drawn: $p_j(s_j|R_i^{t+1})$ equals the number of times that action s_j occurs in the sample R_i^{t+1} divided by k, for each $s_j \in A_j$. Player i then plays an action that is a best response to this predicted mixed strategy: $s_i(t + 1) \in BR_i(R_i^{t+1}) := \arg\max \{\sum_{s_j \in A_j} \left(p_j(s_j|R_i^{t+1}) \cdot u_i(s_i, s_j)\right) \mid s_i \in A_i\}$.

The decision making process described above is called unperturbed adaptive play with memory size m and sample size k. Through an adaptive play process, self-enforcing patterns of play, called conventions, can

[5] The only possible hick-up is that the denominator could equal 0, but that is ruled out when player 1 has two actions that differ to them in more than name only.

[6] Throughout, we use k for the sample size because we already use s for strategies.

emerge.

Definition 1. *A* convention *is a state $h(t|m)$ that entirely consists of m repetitions of the same Nash equilibrium s^* of the game G.*

When a convention is reached in which the Nash equilibrium s^* is played, then the players can only sample the others playing their part of s^* and thus all players have a best response to play their part of s^*. That means that adaptive play predicts that the players can keep playing s^* in all subsequent periods. If the Nash equilibrium s^* is strict, then the best responses are unique and, without perturbations, the players will keep playing s^* indefinitely.

3 Minimally incomplete sampling

In Young [1998], Young proved that in 2×2 coordination games, unperturbed adaptive play will reach a convention as long as sampling is sufficiently incomplete. Incomplete sampling means that the players sample only a faction of the records in memory and in Young [1998] the specific limit for sampling to be "sufficiently" incomplete is $k \leq \frac{m}{2}$, meaning that players sample at most half of all the records available in memory. We relax this bound substantially and show that in 2×2 coordination games, *any* degree of incomplete sampling is sufficient for a convention to eventually be reached.

Lemma 1 will be used in the proof of Theorems 1, 2 and 3.

Lemma 1. *Let G be a 2×2 coordination game and let $s^* = (s_1^*, s_2^*)$ be a (pure-strategy) Nash equilibrium of G. Consider unperturbed adaptive play with memory size m and sample size $k \leq m$. Let $t > m$ be a period in which each player $i \in \{1, 2\}$ can play s_i^* as a best response to their sampled history, so that there is a positive probability that the strategy-tuple s^* is played in period t. Then the convention of playing s^* can be reached with positive probability.*

Proof of Lemma 1. Using induction, we show that there exists a positive probability that s^* is played in periods t through $t + m - 1$, so that the convention of playing s^* is reached.

Base Step: By assumption, the strategy-tuple $s^* = (s_1^*, s_2^*)$ is played with positive probability in period t.

Inductive Step: Let $\hat{t} \geq t$ and suppose that it has already been demonstrated that each player $i \in \{1, 2\}$ can play s_i^* as a best response to their sampled history in period $\hat{t}$, so that there is a positive probability that the strategy-tuple s^* is played in period $\hat{t}$. It will be shown that there is a positive probability that s^* is played in period $\hat{t} + 1$ as part of adaptive play.

For each player $i \in \{1, 2\}$, let $R_i^{\hat{t}}$ be a sampled history of player i in period $\hat{t}$ such that $s_i^* \in BR_i(R_i^{\hat{t}})$, and let $s_i(\hat{t}) = s_i^*$. Then there is a positive probability that each player i draws a sample $R_i^{\hat{t}+1}$ that is obtained by replacing one of the records in $R_i^{\hat{t}}$ with $s_j(\hat{t}) = s_j^*$ $(j \neq i)$. If the replaced record is equal to s_j^*, then this does not change the frequency of s_j^* in i's sample, and if the replaced record is not equal to s_j^*, then this increases the frequency of s_j^* in i's sample. If $s^* = (a_1, b_1)$, then $p_j(s_j^* | R_i^{\hat{t}+1}) \geq p_j(s_j^* | R_i^{\hat{t}}) \geq \alpha_j$, where the last step holds because $s_i^* \in BR_i(R_i^{\hat{t}})$. Similarly, if $s^* = (a_2, b_2)$, then $p_j(s_j^* | R_i^{\hat{t}+1}) \geq p_j(s_j^* | R_i^{\hat{t}}) \geq 1 - \alpha_j$. In both cases, it follows that $s_i^* \in BR_i(R_i^{\hat{t}+1})$.

Therefore, there is a positive probability that s^* is played in period $\hat{t} + 1$ as part of adaptive play.

Conclusion: Using the inductive step $m - 1$ times, it has thus been shown that there exists a positive probability that s^* is played in periods t through $t + m - 1$, so that the convention of playing s^* is reached. ■

Lemma 1 exploits the fact that in a 2×2 game, when player i's Nash equilibrium action s_i^* is a best response to the other player j's mixed strategy, and subsequently, the probability that player j plays s_j^* (weakly) increases, then s_i^* is still a best response by player i. Loosely speaking, it seems fairly intuitive that a when the other player plays their Nash equilibrium action with larger probability, this will increase a player's incentive to play their best response to that action. However, the following example demonstrates that this intuition does not extend to larger games.

Example 1. *Consider the* 3×3 *coordination game*

		Player 2		
		b_1	b_2	b_3
	a_1	$1, 1$	$0, 0$	$0, 0$
Player 1	a_2	$0, 0$	$2, 2$	$-3, -3$
	a_3	$0, 0$	$-3, -3$	$2, 2$

Suppose that $t > m \geq k = 2$, $R_1^t = \{b_3, b_2\}$ *and* $R_2^t = \{a_2, a_3\}$. *Given the distribution of the sampled actions of player 2 in* R_1^t, *player 1 has an expected payoff of* 0 *if they play* a_1, $-\frac{1}{2}$ *if they play* a_2, *and* $-\frac{1}{2}$ *if they play* a_3, *so that* $BR_1(R_1^t) = \{a_1\}$. *Given the distribution of the sampled actions of player 1 in* R_2^t, *player 2 has an expected payoff of* 0 *if they play* b_1, $-\frac{1}{2}$ *if they play* b_2, *and* $-\frac{1}{2}$ *if they play* b_3, *so that* $BR_2(R_2^t) = \{b_1\}$. *Thus, the strict Nash equilibrium* (a_1, b_1) *is played with positive probability period* t.

Let $s(t) = (a_1, b_1)$ and suppose that in period $t+1$ both players sample the record played in period t. Assuming that player 1 draws the record $s_2(t) = b_1$ instead of one of the two records b_2 or b_3 drawn in R_1^t, there are two possibilities, namely $R_1^{t+1} = \{b_1, b_2\}$ and $R_1^{t+1} = \{b_1, b_3\}$. Because $BR_1(\{b_1, b_2\}) = \{a_2\}$ and $BR_1(\{b_1, b_3\}) = \{a_3\}$, it is no longer a best response for player 1 to play b_1 if they replace any of the records that they sampled at time t with the record of player 2 playing b_1 at time t. It is thus possible that adaptive play leads players away from the strict Nash equilibrium (a_1, b_1) after a period in which that strategy profile is played by the two players.

While Example 1 demonstrates that the proof that we provided of Lemma 1 is not valid for coordination games in which players have more than 2 actions, the following example demonstrates that when sampling is complete ($k = m$), the statement of the lemma is not necessarily true for such games.

Example 2. *Consider the game in Example 1 and let $k = m = 2$. Consider a period $t > 2$ such that $h(t \mid m) = \big((a_2, b_3), (a_3, b_2)\big)$. Because sampling is complete, $R_1^{t+1} = \{b_2, b_3\}$ and $R_2^{t+1} = \{a_2, a_3\}$. Because $BR_1(\{b_2, b_3\}) = \{a_1\}$ and $BR_2(\{a_2, a_3\}) = \{b_1\}$, necessarily $s(t+1) = (a_1, b_1)$. Thus, $h(t+1 \mid m) = \big((a_3, b_2), (a_1, b_1)\big)$, $R_1^{t+2} = \{b_1, b_2\}$, and $R_2^{t+2} = \{a_1, a_3\}$. Therefore, $BR_1(R_1^{t+2}) = \{a_2\}$ and $BR_2(R_2^{t+2}) = \{b_3\}$, and necessarily $s(t+2) = (a_2, b_3)$. It follows that in period $t+3$, the players see the history $h(t+2 \mid m) = \big((a_1, b_1), (a_2, b_3)\big)$, so that $R_1^{t+3} = \{b_1, b_3\}$ and $R_2^{t+3} = \{a_1, a_2\}$, and the players' best responses are $BR_1(R_1^{t+3}) = \{a_3\}$ and $BR_2(R_1^{t+3}) = \{b_2\}$. After playing $s(t+3) = (a_3, b_2)$, the history of the last m plays is $h(t+3 \mid m) = \big((a_2, b_3), (a_3, b_2)\big)$ and the adaptive play process has thus returned to the same state it was in during period t.*

We have demonstrated that adaptive play with memory size 2 and complete sampling (sample size 2) will keep cycling from (a_2, b_3) to (a_3, b_2) to (a_1, b_1), to (a_2, b_3), to (a_3, b_2), and so on. Thus, although the strict Nash equilibrium: (a_1, b_1) is played every third period as part of this sequence, the process never reaches the convention of playing (a_1, b_1).

Theorem 1. *Let G be a 2×2 coordination game with Nash equilibria (a_1, b_1) and (a_2, b_2), in which at least one of the two Nash equilibria is strict (i.e., either $u_1(a_1, b_1) > u_1(a_2, b_1)$ and $u_2(a_1, b_1) > u_2(a_1, b_2)$, or $u_1(a_2, b_2) > u_1(a_1, b_2)$ and $u_2(a_2, b_2) > u_2(a_2, b_1)$). From any initial state, unperturbed adaptive play with memory size m and sample size $k < m$ converges with probability one to a convention corresponding to a strict Nash equilibrium and locks in.*

Proof of Theorem 1. In light of Lemma 1, it suffices to demonstrate that there exists a period $t > m$ in which a strict Nash equilibrium $s^* = (s_1^*, s_2^*)$ is played with positive probability, because then the convention of playing s^* can be reached with positive probability, and once that convention is reached, the players will keep playing s^* indefinitely.

Without loss of generality, assume that the Nash equilibrium (a_1, b_1) is strict. Consider unperturbed adaptive play with memory size m and sample size $k < m$ starting from an arbitrary initial state. Consider an arbitrary period $t > m$ and the history $h(t) = (s(1), s(2), s(3), ..., s(t))$ of plays up to and including time t. We distinguish three cases.

Case 1. In period $t+1$ it is possible for the players to draw samples R_i^{t+1}, $i = 1, 2$, such that $a_1 \in BR_1(R_1^{t+1})$ and $b_1 \in BR_2(R_2^{t+1})$. Then there is a positive probability that $s(t+1) = (a_1, b_1)$.

Case 2. In period $t+1$ it is possible for the players to draw samples R_i^{t+1}, $i = 1, 2$, such that $a_2 \in BR_1(R_1^{t+1})$ and $b_2 \in BR_2(R_2^{t+1})$. There is a positive probability that $s(t+1) = (a_2, b_2)$. If the Nash equilibrium (a_2, b_2) is strict, then we have reached a period in which the players play a strict Nash equilibrium.

If the Nash equilibrium (a_2, b_2) is not strict, then $u_1(a_2, b_2) = u_1(a_1, b_2)$ or $u_2(a_2, b_2) = u_2(a_2, b_1)$ (or both). Assume, without loss of generality, that $u_1(a_2, b_2) = u_1(a_1, b_2)$ (and $u_2(a_2, b_2) \geq u_2(a_2, b_1)$). Then $BR_1(R_1^{t+1}) = \{a_1, a_2\}$ and thus $a_1 \in BR_1(R_1^{t+1})$. Thus, $s(t+1) = (a_1, b_2)$ is played with positive probability in the adaptive play process. For the next $k-1$ periods, regardless of the actions that player 2 plays and the samples that player 1 draws, player 1 can keep playing $s_1(\hat{t}) = a_1$, $\hat{t} = t+2, \ldots, t+k$, as a best response. Then in period $t+k+1$, player 2 can draw a sample R_2^{t+k+1} from player 1's actions that consists of k instances of player 1 playing a_1, so that $b_1 \in BR_2(R_2^{t+k+1})$. Thus, there is a positive probability that $s(t+k+1) = (a_1, b_1)$.

Case 3. If in period $t+1$ it is not possible for the players to draw samples R_i^{t+1}, $i = 1, 2$, such that $s_i \in BR_i(R_i^{t+1})$ for $i = 1, 2$ and (s_1, s_2) is a Nash equilibrium of G, then, without loss of generality, assume that $BR_1(R_1^{t+1}) = \{a_1\}$ for all samples that player 1 can draw, and $BR_2(R_2^{t+1}) = \{b_2\}$ for all samples that player 2 can draw, so that $s(t+1) = (a_1, b_2)$.

This implies that in $h(t \mid m)$ player 2 played b_2 at most β_2 times, where β_2 is the largest number in $\{0, 1, \ldots, k-1\}$ that is strictly lower than $(1 - \alpha_2) \times k$.[7] Similarly, in $h(t \mid m)$ player 1 played a_1 at most β_1 times, where β_1 is the largest number in $\{0, 1, \ldots, k-1\}$ that is strictly lower than $\alpha_1 \times k$.[8] However, $s(t+1) = (a_1, b_2)$, so that the number of times that player 1 (resp. 2) plays action a_1 (resp. b_2) in $h(t+1 \mid m)$

[7] We remind the reader that α_2 is the probability such that action a_1 is a best response by player 1 to a mixed strategy p_2 of player 2 if and only if $p_2(b_1) \geq \alpha_2$. Also, because (a_1, b_1) is a strict Nash equilibrium, $\alpha_2 < 1$, so that $(1 - \alpha_2) \times k > 0$.

[8] We remind the reader that α_1 is the probability such that action b_1 is a best response by player 2 to a mixed strategy p_1 of player 1 if and only if $p_1(a_1) \geq \alpha_1$. Note that $\alpha_1 > 0$, because otherwise $b_1 \in BR_2(R_2^{t+1})$ regardless of the sample that player 2 draws.

is either equal to that in $h(t \mid m)$ (in case $s_1(t-m+1) = a_1$, resp. $s_2(t-m+1) = b_2$) or one higher. As long as these numbers do not exceed β_1, resp. β_2, the players will keep playing $s(\hat{t}) = (a_1, b_2)$ in periods $\hat{t} \geq t+2$. This clearly cannot persist because after m periods the players would only have plays (a_1, b_2) in recent memory.

Let $\hat{t} \geq t+1$ be the first period in which either player 1 played a_1 more than β_1 times in $h(\hat{t} \mid m)$ or player 2 played b_2 more than β_2 times in $h(\hat{t} \mid m)$ (or both). Without loss of generality, assume that player 1 played β_1 instances of a_1 in $h(\hat{t}-1 \mid m)$ and β_1+1 instances of a_1 in $h(\hat{t} \mid m)$. Then in period $\hat{t}+1$, player 2 can draw a sample $R_2^{\hat{t}+1}$ that contains β_1+1 instances of player 1 playing a_1, and play $s_2(\hat{t}+1) = b_1 \in BR_2(R_2^{\hat{t}+1})$. Also, player 2 played at most β_2 instances of b_2 in $h(\hat{t}-1 \mid m)$), and thus at most β_2+1 instances of b_2 in $h(\hat{t} \mid m)$). Thus, because $k < m$, in period $\hat{t}+1$, player 1 can draw a sample $R_1^{\hat{t}+1}$ that contains no more than β_2 instances of player 2 playing b_2, and play $s_1(\hat{t}+1) = a_1 \in BR_1(R_1^{\hat{t}+1})$. Thus, there is a positive probability that $s(\hat{t}+1) = (a_1, b_1)$.

Conclusion. The three cases we considered are exhaustive and thus we have shown that, starting from any period $t > m$ and with any history of play at that time, we can find a period in which there is a positive probability that the players play a strict Nash equilibrium in the adaptive play process with sample size $k < m$. Lemma 1 then establishes that the convention of playing that strict Nash equilibrium can be reached with positive probability, and then the process is locked in. ■

Note that in the proof of Theorem 1, there is only one instance in which we use that sampling is incomplete ($k < m$), and that is in Case 3, where we need it to guarantee that it cannot be the case that the adaptive play process can get "stuck" in a situation where both players mis-coordinate in every period, oscillating between (a_1, b_2) and (a_2, b_1) and necessarily switching actions in exactly the same periods. If the game and sample sizes are such that this cannot happen anyway, then we do not need sampling to be incomplete at all, and we can have $k = m$. We use the notation $\lceil \cdot \rceil$ to denote the rounding up of any real number to the smallest natural number that is at least as large.[9]

Theorem 2. *Let G be a 2×2 coordination game with Nash equilibria (a_1, b_1) and (a_2, b_2), in which at least one of the two Nash equilibria is strict (i.e., either $u_1(a_1, b_1) > u_1(a_2, b_1)$ and $u_2(a_1, b_1) > u_2(a_1, b_2)$, or $u_1(a_2, b_2) > u_1(a_1, b_2)$ and $u_2(a_2, b_2) > u_2(a_2, b_1)$) and such that $\alpha_1 \neq 1 - \alpha_2$.*[10] *Let the sample size k be such that $\lceil \alpha_1 \times k \rceil \neq \lceil (1-\alpha_2) \times k \rceil$ or $\lceil \alpha_2 \times k \rceil \neq \lceil (1-\alpha_1) \times k \rceil$.*[11,12] *From any initial state, unperturbed*

[9] So, if n is a natural number itself, then $\lceil n \rceil = n$. Also, we include 0 in the set of natural numbers.

[10] Thus, the smallest probability for player 1 to play a_1 such that action b_1 is a best response by player 2 is not equal to the smallest probability for player 2 to play b_2 such that action a_2 is a best response by player 1.

[11] Note that if the game G is such that α_1 and $1-\alpha_2$ are close, then this will require a large sample size.

[12] $\lceil \alpha_1 \times k \rceil \neq \lceil (1-\alpha_2) \times k \rceil$ does not necessarily imply $\lceil \alpha_2 \times k \rceil \neq \lceil (1-\alpha_1) \times k \rceil$. An example of this can be found in

adaptive play with memory size m and sample size $k \leq m$ converges with probability one to a convention corresponding to a strict Nash equilibrium and locks in.

Proof of Theorem 2. If $k < m$, then Theorem 1 applies. So, suppose that $k = m$, i.e, sampling is complete in the sense that players see *all* of the past m records.

Consider an adaptive play process with $k = m$. If in some period $t > m$ the players coordinate, i.e., $s(t) = (a_1, b_1)$ or $s(t) = (a_2, b_2)$, then we can apply cases 1 or 2 in the proof of Theorem 1 to establish that there is a positive probability that the players play a strict Nash equilibrium (note that these cases do not depend on $k < m$). Lemma 1 then establishes that the convention of playing that strict Nash equilibrium can be reached with positive probability, and then the process is locked in.

Thus, it remains to consider the possibility that the players mis-coordinate in all periods, i.e., $s(t) \in \{(a_1, b_2), (a_2, b_1)\}$ for all t. We will demonstrate that this cannot happen because one of $\lceil \alpha_1 \times k \rceil \neq \lceil (1 - \alpha_2) \times k \rceil$ or $\lceil \alpha_2 \times k \rceil \neq \lceil (1 - \alpha_1) \times k \rceil$ implies that an adaptive play process with $k = m$ cannot result in string of mis-coordinated plays $s(1), s(2), \ldots$ with $s(t) \in \{(a_1, b_2), (a_2, b_1)\}$ for all t. Without loss of generality assume $\lceil \alpha_1 \times m \rceil \neq \lceil (1 - \alpha_2) \times m \rceil$. If, in some period $t > m$,[13] the players observe a history of play that consists of a string of m instances of (a_1, b_2) having been played in the previous m periods, player 2's unique best response is to play b_1 in the next period or player 1's unique best response is to play a_2 in the next period.[14] Thus, any string of mis-coordinated plays that contains a string of more than m subsequent plays of (a_1, b_2) cannot be the result of an adaptive play process. Similarly, any string of mis-coordinated plays that contains a string of more than m subsequent plays of (a_2, b_1) cannot be the result of an adaptive play process. We conclude that if the players mis-coordinate in all periods, and they follow an adaptive play process, then the process needs to switch repeatedly between playing (a_1, b_2) and (a_2, b_1).

For player 1 to switch to playing a_2, they need to observe $\lceil (1 - \alpha_2) \times m \rceil$ instances of player 2 playing b_2, and for player 2 to switch to playing b_1, they need to observe $\lceil \alpha_1 \times m \rceil$ instances of player 1 playing a_1. However, in all periods $t > m$, because $k = m$, player 1 samples as many records of player 2 playing b_2 as player 2 samples records of player 1 playing a_1. Thus, $\lceil \alpha_1 \times m \rceil \neq \lceil (1 - \alpha_2) \times m \rceil$ implies that the players will not switch from playing (a_1, b_2) to playing (a_2, b_1) in the same period when they follow an adaptive play process. ■

Example 4.

[13] We consider only periods $t > m$ to ensure that the players have m periods' plays available in memory.

[14] This uses $\lceil \alpha_1 \times m \rceil \neq \lceil (1 - \alpha_2) \times m \rceil$, which implies that it cannot be the case that player 1 can best respond by playing a_1 *and* player 2 can best respond by playing b_2 after both observe m instances of (a_1, b_2) having been played.

4 Perturbed adaptive play in 2×2 games

Now consider the adaptive play process as modeled in Section 2 but where players have a small probability of playing an action that is not their best response. Specifically, in every round players now play a strategy at random with probability ε and with probability $1-\varepsilon$ they play a best response to their drawn sample, R_i^t. As such, an action that is not in the set of possible best responses to samples drawn from the memory can be played with probability $\varepsilon/2$.[15] We shall refer to such actions as *mistakes*. This process is called perturbed adaptive play. Allowing for mistakes makes transitions possible between conventions, even those in which a strict Nash equilibrium is played. Continuing with the setup established in Section 2 where (a_1, b_1) and (a_2, b_2) are Nash equilibria, denote by h_i the convention corresponding to (a_i, b_i), i.e., the state that consists of m repetitions of (a_i, b_i). Now consider the transition from h_i to h_j. Let the resistance, denoted $r_{i,j}^{k,m}$, be the minimum number of mistakes necessary to make the transition from h_i to h_j in the perturbed adaptive play process. Young [1998] shows that the resistance between conventions is independent of m when $k \leq m/2$. Specifically, the resistance of moving from h_2 to h_1 equals $\min\left(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil\right)$ and the resistance of moving from h_1 to h_2 equals $\min\left(\lceil (1-\alpha_1) \times k \rceil, \lceil (1-\alpha_2) \times k \rceil\right)$. However, we demonstrate below that when sampling is less incomplete ($k > m/2$), the resistances may be larger and depend on m.

Theorem 3. *Let G be a 2×2 coordination game with Nash equilibria (a_1, b_1) and (a_2, b_2). Consider unperturbed adaptive play with memory size m and sample size $k \leq m$. The resistance of moving from h_2 to h_1 equals $r_{2,1}^{k,m} = \min\left(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil\right) + \max\left(\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil - m, 0\right)$ and the resistance of moving from h_1 to h_2 equals $r_{1,2}^{k,m} = \min\left(\lceil (1-\alpha_1) \times k \rceil, \lceil (1-\alpha_2) \times k \rceil\right) + \max\left(\lceil (1-\alpha_1) \times k \rceil + \lceil (1-\alpha_2) \times k \rceil - m, 0\right)$.*

Proof of Theorem 3. We compute the resistance $r_{2,1}^{k,m}$. Similarly to Case 2 of the proof of Theorem 1, we derive that no mistakes are necessary to move from h_2 to h_1 if equilibrium (a_2, b_2) is not strict. In that case, either $\alpha_1 = 0$ or $\alpha_2 = 0$ or both hold and the expression for $r_{2,1}^{k,m}$ in the statement of the theorem indeed produces a resistance equal to 0.

So assume that equilibrium (a_2, b_2) is strict and let t be a period such that $h(t|m) = h_2$, i.e., the system is in convention h_2. Because (a_2, b_2) is strict, (a_2, b_2) will continue to be played if the players do not make any mistakes. To reach convention h_1, it is necessary to reach a period in which both a_1 and b_1 can be played as best responses to samples drawn by the players.[16] Reaching a period in which both a_1 and b_1 can be played as best responses to samples drawn by the players is also a sufficient condition for the process to

[15] The 1/2 comes from the fact that each player has two actions, each of which they choose with equal probability when they play an action at random.

[16] Note that this is a property that is satisfied by convention h_1.

reach convention h_1 without further mistakes (see Lemma 1). Thus, starting from convention h_2, we need to determine the minimum number of mistakes (which will be positive) necessary to build a length-m history of play from which both players can draw samples of size k such that a_1 and b_1 are best responses. For this condition to be met in some period T, player 1 must have played a_1 at least $\lceil \alpha_1 \times k \rceil$ times and player 2 must have played b_1 at least $\lceil \alpha_2 \times k \rceil$ times in periods $T-m$ through $T-1$. Clearly, this can be accomplished by having player 1 make a mistake and play a_1 a total of $\lceil \alpha_1 \times k \rceil$ times *and* having player 2 make a mistake and play b_1 a total of $\lceil \alpha_2 \times k \rceil$ times in $\max\left(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil\right)$ consecutive periods. This gives an upper bound of $\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil$ for $r_{2,1}^{k,m}$.

The number of mistakes can be lowered by decreasing the number of periods in which both players make a mistake, so that players can sample each other's mistakes and potentially play a_1 and/or b_1 as best responses. At the extreme, when sample sizes are sufficiently incomplete so that players can keep sampling mistakes long enough, it suffices for one player to make enough mistakes to make their action in (a_1, b_1) a best response by the other player, and we obtain the lower bound $\min\left(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil\right)$ for $r_{2,1}^{k,m}$. We consider this case first.

Case 1. $\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil \leq m$.[17]

Starting in period $t+1$, suppose player 1 makes $\lceil \alpha_1 \times k \rceil$ consecutive mistakes and plays a_1 in periods $t+1, \ldots, t+\lceil \alpha_1 \times k \rceil$. During each of these periods, player 2 can sample no more than $\lceil \alpha_1 \times k \rceil - 1$ instances of player 1 playing a_1 and can only play b_2 as a best response.

Because $\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil \leq m$, in each of the periods $t+\lceil \alpha_1 \times k \rceil + 1$ through $t+\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil$, player 2 can sample all $\lceil \alpha_1 \times k \rceil$ instances of a_1 that player 1 played in periods $t+1$ to $t+\lceil \alpha_1 \times k \rceil$ and play b_1 as a best response. In these periods, player 1 can sample no more than $\lceil \alpha_2 \times k \rceil - 1$ plays of b_1 and can only play a_2 as a best response.

	Play										
Period	$t-m+1$	$\ldots$	t	$t+1$	$\ldots$	$t+\lceil \alpha_1 \times k \rceil$	$t+\lceil \alpha_1 \times k \rceil+1$	$\ldots$	$t+\lceil \alpha_1 \times k \rceil+\lceil \alpha_2 \times k \rceil$	$t+\lceil \alpha_1 \times k \rceil+\lceil \alpha_2 \times k \rceil+1$	$\ldots$
Player 1	a_2	a_2	a_2	a_1	a_1	a_1	a_2	a_2	a_2	a_1	a_1
Player 2	b_2	b_2	b_2	b_2	b_2	b_2	b_1	b_1	b_1	b_1	b_1

The color red denotes actions which necessarily are mistakes. Actions colored blue can be played as a best response.

Because $\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil \leq m$, in period $t+\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil + 1$ it is possible for player 2 to sample all $\lceil \alpha_1 \times k \rceil$ player 1's plays of a_1 in periods $t+1$ through $t+\lceil \alpha_1 \times k \rceil$, while player 1 samples all $\lceil \alpha_2 \times k \rceil$ player 2's plays of b_1 in periods $t+\lceil \alpha_1 \times k \rceil + 1$ through $t+\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil$. Thus, both a_1 and b_1 can be played as best responses by the players in period $t+\lceil \alpha_1 \times k \rceil + \lceil \alpha_2 \times k \rceil + 1$ and the process can reach convention h_1 without further mistakes (see Lemma 1).

[17] In this case, derivations are similar to those in Young [1998].

The process we just described reaches convention h_1 from convention h_2 with exactly $\lceil\alpha_1 \times k\rceil$ mistakes by starting in period $t+1$ with player 1 making $\lceil\alpha_1 \times k\rceil$ consecutive mistakes and playing a_1 in periods $t+1,\ldots,t+\lceil\alpha_1\times k\rceil$. If instead we start in period $t+1$ with player 2 making $\lceil\alpha_2\times k\rceil$ consecutive mistakes and playing b_1 in periods $t+1,\ldots,t+\lceil\alpha_2\times k\rceil$, we obtain a process that reaches convention h_1 from convention h_2 with exactly $\lceil\alpha_2 \times k\rceil$ mistakes.

Because either player 1 must have played a_1 at least $\lceil\alpha_1 \times k\rceil$ times to allow player 2 to play b_1 as a best response, or player 2 must have played b_1 at least $\lceil\alpha_2 \times k\rceil$ times to allow player 1 to play a_1 as a best response, the minimum number of mistakes necessary to reach convention h_1 from convention h_2 equals $\min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right)$. Since (at least) one of the two processes we described reaches convention h_1 from convention h_2 with exactly $\min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right)$ mistakes, we have demonstrated that $r_{2,1}^{k,m} = \min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) = \min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) + \max\left(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0\right)$.[18]

Case 2. $\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil > m$.

In order to make the transition from h_2 to h_1, at least $\lceil\alpha_1 \times k\rceil$ plays of b_1 *and* $\lceil\alpha_2 \times k\rceil$ plays of a_1 must occur within m periods. However, $\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil > m$ implies that in order to achieve this condition, a_1 and b_1 must be played in the same period a minimum of $\ell := \lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m$ times.[19] Because player 1 cannot play a_1 as a best response until player 2 has played b_1 at least $\lceil\alpha_2 \times k\rceil$ times, and player 2 cannot play b_1 as a best response until player 1 has played a_1 at least $\lceil\alpha_1 \times k\rceil$ times, the ℓ concurrent plays of a_1 and b_1 require $2\times\ell$ mistakes. Thus, in this case, we need at least an additional ℓ mistakes compared to Case 1. We demonstrate that we do not need more than an additional ℓ mistakes by describing a transition from h_2 to h_1 with exactly $\min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) + \max\left(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0\right)$ mistakes.

Starting in period $t+1$, suppose player 1 makes $\lceil\alpha_1 \times k\rceil$ consecutive mistakes and plays a_1 in periods $t+1,\ldots,t+\lceil\alpha_1\times k\rceil$. During each of these periods, player 2 can sample no more than $\lceil\alpha_1\times k\rceil - 1$ instances of player 1 playing a_1 and can only play b_2 as a best response. Suppose that in the last ℓ of these periods, $t+m-\lceil\alpha_2 \times k\rceil+1$ through $t+\lceil\alpha_1 \times k\rceil$, player 2 makes ℓ consecutive mistakes and plays b_1.

	Play													
Period	$t-m+1$	$\ldots$	t	$t+1$	$\ldots$	$t+m-\lceil\alpha_2\times k\rceil$	$t+m-\lceil\alpha_2\times k\rceil+1$	$\ldots$	$t+\lceil\alpha_1\times k\rceil$	$t+\lceil\alpha_1\times k\rceil+1$	$\ldots$	$t+m$	$t+m+1$	$\ldots$
Player 1	a_2	a_2	a_2	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_1	a_1
Player 2	b_2	b_2	b_2	b_2	b_2	b_2	b_1	b_1	b_1	b_1	b_1	b_1	b_1	b_1

The color red denotes actions which necessarily are mistakes. Actions colored blue can be played as a best response.

If $\lceil\alpha_1 \times k\rceil = \lceil\alpha_2 \times k\rceil = m$, then $\ell = m$ and the process described so far has reached convention h_1 with $2\times m$ mistakes and this convention cannot be reached from h_2 with fewer mistakes. Thus, $r_{2,1}^{k,m} = 2\times m =$

[18]Note that when $k \le m/2$, the condition $\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil \le m$ is satisfied regardless of the values of α_1 and α_2. This is why in Young [1998] the resistance $r_{2,1}^{k,m}$ is given as $r_{2,1}^{k} = \min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right)$.

[19]Note that $\ell = \lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m \le m$.

$\min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) + \max\left(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0\right)$.

It remains to consider the case when $\lceil\alpha_1 \times k\rceil < m$. In that case, in periods $t + \lceil\alpha_1 \times k\rceil + 1$ through $t + m$, player 2 can sample all $\lceil\alpha_1 \times k\rceil$ of player 1's plays of a_1 in periods $t + 1$ through $t + \lceil\alpha_1 \times k\rceil$, and play b_1 as a best response. Because, by definition of ℓ, $m = \left(\lceil\alpha_1 \times k\rceil - \ell\right) + \ell + \left(\lceil\alpha_2 \times k\rceil - \ell\right)$, in these periods, player 1 can sample no more than $\lceil\alpha_2 \times k\rceil - 1$ plays of b_1 and can only play a_2 as a best response.

In period $t + m + 1$, it is possible for player 1 to sample all $\lceil\alpha_2 \times k\rceil$ player 2's plays of b_1 in periods $t + m - \lceil\alpha_2 \times k\rceil + 1$ through $t + m$, while player 2 samples all $\lceil\alpha_1 \times k\rceil$ player 1's plays of a_1 in periods $t + 1$ through $t + \lceil\alpha_1 \times k\rceil$. Thus, both a_1 and b_1 can be played as best responses by the players in period $t + m + 1$ and the process can reach convention h_1 without further mistakes (see Lemma 1).

The process we just described reaches convention h_1 from convention h_2 with exactly $\lceil\alpha_1 \times k\rceil + \ell$ mistakes if $\alpha_1 \leq \alpha_2$. Analogously, we can describe a process that reaches convention h_1 from convention h_2 with exactly $\lceil\alpha_2 \times k\rceil + \ell$ mistakes if $\alpha_2 \leq \alpha_1$. Thus, we have identified a process that reaches convention h_1 from convention h_2 with exactly the minimum number of mistakes $\min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) + \ell$ that we identified as necessary, and we have demonstrated that $r_{2,1}^{k,m} = \min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) + \max\left(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0\right)$.

Conclusion. We demonstrated that $r_{2,1}^{k,m} = \min\left(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\right) + \max\left(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0\right)$. The resistance $r_{1,2}^{k,m}$ is now easily obtained by using $1 - \alpha_1$ and $1 - \alpha_2$ instead of α_1 and α_2, resulting in $r_{1,2}^{k,m} = \min\left(\lceil(1 - \alpha_1) \times k\rceil, \lceil(1 - \alpha_2) \times k\rceil\right) + \max\left(\lceil(1 - \alpha_1) \times k\rceil + \lceil(1 - \alpha_2) \times k\rceil - m, 0\right)$. ■

The interest in resistances of moving between conventions stems from the fact that, when the probability of making mistakes (i.e., the degree of perturbation in the process) becomes vanishingly small, the perturbed adaptive play process converges on the conventions that are hardest to leave and easiest to reach.

In a game with exactly two Nash equilibria, and thus two conventions, a convention is stochastically stable if and only if the resistance in the transition away from it is at least as large as the resistance in the transition towards it (see Young [1993]). In the model studied in Young [1993], sampling is sufficiently incomplete ($k \leq m/2$) for the resistances between conventions to be independent of m., so that stochastic stability of conventions is also independent of m. In contrast, as we demonstrated in Theorem 3, the resistances may be larger and depend on m when sampling is less incomplete ($k > m/2$). This opens up the possibility that the degree of incomplete sampling influences which conventions are stochastically stable. We turn to this next.

The following example shows that changing k when m is fixed may change which states are stochastically stable.

Example 3. *Consider the 2×2 coordination game*

		Player 2	
		b_1	b_2
Player 1	a_1	$10, 13$	$0, 0$
	a_2	$2, 3$	$12, 10$

In this game $\alpha_1 = 7/20$ and $\alpha_2 = 3/5$. Suppose $m = 10$ and k increases from 5 to 10. When $k = 5$, $\lceil \alpha_1 \times k \rceil = \lceil (1 - \alpha_2) \times k \rceil = 2$, $\lceil (1 - \alpha_1) \times k \rceil = 4$ and $\lceil \alpha_2 \times k \rceil = 3$. So $r_{2,1}^{5,10} = 2$ and $r_{1,2}^{5,10} = 2$, which means both conventions h_1 and h_2 are stochastically stable. However, when $k = 10$, $\lceil \alpha_1 \times k \rceil = \lceil (1 - \alpha_2) \times k \rceil = 4$, $\lceil (1 - \alpha_1) \times k \rceil = 7$ and $\lceil \alpha_2 \times k \rceil = 6$. So $r_{2,1}^{10,10} = 4$ and $r_{1,2}^{10,10} = 5$ which means in this case only h_1 is stochastically stable.

Example 3 shows that changing the degree of incomplete sampling, k/m, by varying k and keeping m fixed can change which states are stochastically stable through the added $max(\cdot)$ component of the resistance calculation.[20] However, stochastic stability does not change when the degree of incomplete sampling, k/m, changes by varying m while keeping k fixed. Next, we prove that when sampling is incomplete, changing m alone does not change which conventions are stochastically stable.

Theorem 4. *Let G be a $2{\times}2$ coordination game with Nash equilibria (a_1, b_1) and (a_2, b_2). Consider unperturbed adaptive play with fixed sample size k and memory size m such that sampling is incomplete: $m > k$. Stochastic stability of conventions does not depend on memory size m.*

Proof of Theorem 4. Consider the pairwise resistances between h_2 and h_1. We will consider both the case where $\min(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil) \neq \min(\lceil (1-\alpha_1) \times k \rceil, \lceil (1-\alpha_2) \times k \rceil)$ and where $\min(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil) = \min(\lceil (1-\alpha_1) \times k \rceil, \lceil (1-\alpha_2) \times k \rceil)$. We show that when k is held fixed, changing m to some $m > k$ does not affect the comparison between $r_{1,2}^{k,m}$ and $r_{2,1}^{k,m}$.

Case 1. $\min(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil) \neq \min(\lceil (1-\alpha_1) \times k \rceil, \lceil (1-\alpha_2) \times k \rceil)$.[21]
Without loss of generality assume $\min(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil) < \min(\lceil (1-\alpha_1) \times k \rceil, \lceil (1-\alpha_2) \times k \rceil)$. This means that for all $m \geq 2k$, $r_{2,1}^{k,m} < r_{1,2}^{k,m}$. We will show that when $m < 2k$, the inequality $r_{2,1}^{k,m} < r_{1,2}^{k,m}$ is maintained.

Without loss of generality assume $\alpha_1 \geq \alpha_2$. So, $\min(\lceil \alpha_1 \times k \rceil, \lceil \alpha_2 \times k \rceil) = \lceil \alpha_2 \times k \rceil$ and $\min(\lceil (1 - \alpha_1) \times k \rceil, \lceil (1 - \alpha_2) \times k \rceil) = \lceil (1 - \alpha_1) \times k \rceil$ which means $\lceil \alpha_2 \times k \rceil < \lceil (1 - \alpha_1) \times k \rceil$. It follows that $\alpha_2 < 1 - \alpha_1$ so $\alpha_1 < 1 - \alpha_2$ which means $\lceil \alpha_1 \times k \rceil \leq \lceil (1-\alpha_2) \times k \rceil$. Combining $\lceil \alpha_2 \times k \rceil < \lceil (1-\alpha_1) \times k \rceil$ and

[20] Note that it is already known that stochastic stability may change with k due to the ceiling functions even when $k \leq m/2$.
[21] Note that the restriction $m > k$ is not leveraged in this case.

$\lceil\alpha_1\times k\rceil \leq \lceil(1-\alpha_2)\times k\rceil$ we get $\lceil\alpha_1\times k\rceil+\lceil\alpha_2\times k\rceil < \lceil(1-\alpha_1)\times k\rceil+\lceil(1-\alpha_2)\times k\rceil$. This inequality combined with the fact, due to the fact that $\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil + \lceil(1-\alpha_1)\times k\rceil + \lceil(1-\alpha_2)\times k\rceil \in \{2k, 2k+1, 2k+2\}$ means that $\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil \leq k < \lceil(1-\alpha_1)\times k\rceil + \lceil(1-\alpha_2)\times k\rceil$ as consequence of the fact that the ceiling function can only yield integers. This means that $\max\big(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0\big) = 0 \leq \max\big(\lceil(1-\alpha_1)\times k\rceil + \lceil(1-\alpha_2)\times k\rceil - m, 0\big)$. This result shows that the $\max(\cdot)$ component of resistance in one direction can only be positive if the $\min(\cdot)$ component is strictly larger than the $\min(\cdot)$ component in the opposite direction. Therefore, if $\min(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil) < \min(\lceil(1-\alpha_1)\times k\rceil, \lceil(1-\alpha_2)\times k\rceil)$ then $r_{1,2} < r_{2,1}$ for all $m \geq k$.

So we have shown that the comparison between resistances is unchanged by the size of m in this case.

Case 2. $\min(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil) = \min(\lceil(1-\alpha_1)\times k\rceil, \lceil(1-\alpha_2)\times k\rceil)$.
This means that when $m \geq 2k$, the resistance is equal in both directions: $r_{1,2}^{k,m} = r_{2,1}^{k,m}$. We will show that when $k < m < 2k$, the relationship $r_{1,2}^{k,m} = r_{2,1}^{k,m}$ is maintained. Without loss of generality assume $\alpha_1 \geq \alpha_2$. So, $\min(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil) = \lceil\alpha_2 \times k\rceil$ and $\min(\lceil(1-\alpha_1)\times k\rceil, \lceil(1-\alpha_2)\times k\rceil) = \lceil(1-\alpha_1)\times k\rceil$ which means $\lceil\alpha_2 \times k\rceil = \lceil(1-\alpha_1)\times k\rceil$. It is easily verified that $\lceil\alpha_1 \times k\rceil + \lceil(1-\alpha_1)\times k\rceil \in \{k, k+1\}$ and so $\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil \leq k+1$. Since both m and k are natural numbers and $k < m$ it follows that $k+1 \leq m$. Thus, $\max(\lceil\alpha_1 \times k\rceil + \lceil\alpha_2 \times k\rceil - m, 0) = 0$.

Likewise, $\lceil\alpha_2\times k\rceil+\lceil(1-\alpha_2)\times k\rceil \in \{k, k+1\}$. Because $\lceil\alpha_2\times k\rceil = \lceil(1-\alpha_1)\times k\rceil$, it follows that $\lceil(1-\alpha_1)\times k\rceil+\lceil(1-\alpha_2)\times k\rceil \leq k+1$. Since $k+1 \leq m$, it follows that $\max(\lceil(1-\alpha_1)\times k\rceil+\lceil(1-\alpha_2)\times k\rceil-m, 0) = 0$. Those results combined with the assumption $\min\big(\lceil\alpha_1 \times k\rceil, \lceil\alpha_2 \times k\rceil\big) = \min\big(\lceil(1-\alpha_1)\times k\rceil, \lceil(1-\alpha_2)\times k\rceil\big)$ mean that $r_{1,2}^{k,m} = r_{2,1}^{k,m}$ for all $m > k$ in this case.

Conclusion. The two cases are exhaustive. We have shown that all pairwise resistance comparisons remain unchanged for all $m > k$ given k is fixed. So, it follows that which states are stochastically stable also remains unchanged. ■

Note that we use that sampling is incomplete ($m > k$) only in Case 2 in the proof of Theorem 4. This case covers instances where both conventions h_1 and h_2 are stochastically stable when m is large ($m \geq 2k$). If only one of the conventions is stochastically stable for large m as in Case 1, then we do not need any incomplete sampling to obtain the result that stochastic stability of conventions does not depend on memory size m. However, the following example shows that when both conventions are stochastically stable for large m, decreasing memory size to $m = k$ may render one of the conventions no longer stochastically stable.

Example 4. *Consider the 2×2 coordination game*

		Player 2	
		b_1	b_2
Player 1	a_1	$10, 11$	$0, 0$
	a_2	$0, 1$	$10, 10$

Suppose that $k = 10$ and consider the resistances between h_1 and h_2. Note that in this game $\alpha_1 = 9/20$ and $\alpha_2 = 1/2$. We compute $\lceil \alpha_1 \times k \rceil = \lceil \alpha_2 \times k \rceil = \lceil (1 - \alpha_2) \times k \rceil = 5$ and $\lceil (1 - \alpha_1) \times k \rceil = 6$.

When $m > k = 10$, the resistances between h_1 and h_2 are $r_{1,2}^{k,m} = r_{2,1}^{k,m} = 5$, and both conventions h_1 and h_2 are stochastically stable.

However, when $m = k = 10$, the resistance of moving from h_1 to h_2 increases to $r_{1,2}^{k,m} = 6$ while the resistance of moving from h_2 to h_1 remains unchanged at $r_{2,1}^{k,m} = 5$. Thus, only convention h_1 is stochastically stable.

5 Conclusion

Young's model of adaptive play has been studied and applied to a wide variety of games. However, the boundary of precisely how incomplete sampling needs to be in order for foundational results like convergence to a convention remain unaddressed. We examined the most foundational game, the 2×2 coordination game, and proved that in this case, *any* degree of incomplete sampling is sufficient for the unperturbed adaptive play process to converge to a convention. In addition, we show that in all but some 2×2 games that the criterion of incomplete sampling is unnecessary. We also examine how allowing for minimally incomplete sampling affects the perturbed adaptive process. We identified the function for resistance that is robust to all degrees of sampling, $k \leq m$, and found that increasing the sample size beyond $k/m \leq 1/2$ may result in increased resistance between conventions. However, we also showed that even if the resistance does increase due to the relative size of a fixed sample to a changing memory that the stochastically stable states remain unchanged if sampling is incomplete ($k < m$).

References

K. Basu and J. W. Weibull. Strategy subsets closed under rational behavior. *Economics Letters*, 36(2): 141–146, 1991.

J. Durieu, P. Solal, and O. Tercieux. Adaptive learning and p-best response sets. *International Journal of Game Theory*, 40:735–747, 2011.

G. Ellison. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *The Review of Economic Studies*, 67(1):17–45, 2000.

M. Jensen, B. Sloth, and H. J. Whitta-Jacobsen. The evolution of conventions under incomplete information. *Economic Theory*, 25:171–185, 2005.

A. Khan and R. Peeters. Cognitive hierarchies in adaptive play. *International Journal of Game Theory*, 43: 903–924, 2014.

T. Maruta. On the relationship between risk-dominance and stochastic stability. *Games and Economic Behavior*, 19(2):221–234, 1997.

A. Matros. Clever agents in adaptive learning. *Journal of Economic Theory*, 111(1):110–124, 2003.

S. Morris, R. Rob, and H. S. Shin. p-dominance and belief potential. *Econometrica: Journal of the Econometric Society*, pages 145–157, 1995.

R. Nagel. Unraveling in guessing games: An experimental study. *The American economic review*, 85(5): 1313–1326, 1995.

M. Sáez-Martí and J. W. Weibull. Clever agents in young's evolutionary bargaining model. *Journal of Economic Theory*, 86(2):268–279, 1999.

O. Tercieux. p-best response set. *Journal of Economic Theory*, 131(1):45–70, 2006.

C. Wallace and H. P. Young. Chapter 6 - stochastic evolutionary game dynamics. volume 4 of *Handbook of Game Theory with Economic Applications*, pages 327–380. Elsevier, 2015. doi: https://doi.org/10.1016/B978-0-444-53766-9.00006-9. URL https://www.sciencedirect.com/science/article/pii/B9780444537669000069.

H. P. Young. The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pages 57–84, 1993.

H. P. Young. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, 1998. ISBN 9780691086873. URL http://www.jstor.org/stable/j.ctv10h9d35.